Exploring Network Effects of Point-to-Point Networks: An Investigation of the Spatial Entry Patterns of Southwest Airlines

Jia Yan
Washington State University, Email: jiay@wsu.edu

Xiaowen Fu
Hong Kong Polytechnic University, Email: lgtxfu@polyu.edu.hk

Tae Oum
University of British Columbia, and Embry Riddle University,
Email: tae.oum@sauder.ubc.ca

Acknowledgement: We would like to thank seminar participants at the workshops at Washington State University, and 2008 ASSA-TPUG Conference in New Orleans for helpful comments on an earlier version of the paper. The second author thanks the financial support from the Hong Kong RGC Public Policy Research Grant (PolyU5002-PPR-5).
Abstract

This paper explores network effects in Point-to-Point airline networks by examining the spatial entry patterns of Southwest airlines during the 1990-2006 period. Estimation results from a spatial probit model reveal clear spatial dependence in profitability across different routes served by the carrier. Detailed investigation suggests two main sources of network effects, namely: (1) airport and regional presence, and (2) substitutability of markets. Findings of the paper suggest also that the network effects embedded in Southwest’s Point-to-Point network have many distinguishing features as compared to those identified in a typical Hub-and-Spoke network. This study brings some fresh insights on airline network effects in general, as well as explains the pattern of aggressive network expansions of LCCs in particular.
1. Introduction

It is well known that the Hub-and-Spoke (HS) network, with which travelers can make connecting stops at hub airports before flying to their final destinations, allows airlines to achieve cost savings through “economies of traffic density”\(^1\), to increase flight frequency and thus service quality\(^2\), and to price and compete more strategically\(^3\). Such effects are generally termed as “network effects” in the literature.

Although HS network is the predominant network model in the airline industry, some of the most successful low cost carriers (LCCs)\(^4\) including Southwest Airlines and Jet Blue choose to operate point-to-point networks (PoP) which are comprised of many linear segments without extensive connection operations. In this paper we identify possible network effects of PoP networks by examining the spatial entry patterns of Southwest Airlines, which is the largest LCC in the US domestic market.

Theoretical models in Hendricks et al. (1995 and 1999) suggest that the PoP network allows airlines to compete on price less aggressively. Their models along with the one in Lederer and Nambimadom (1998) predict that airlines may prefer PoP networks if: (1) the sizes of the cities are

\[^1\] Caves et al. (1984), Brueckner and Spiller (1994). Traffic density is calculated by dividing the total traffic volume by the carrier’s network size. Network size is usually defined as the number of origin-destination pairs served by the carrier, or the number of nodes connected in its network.

\[^2\] Morrison and Winston (1987), Berechman and Shy (1998), Brueckner and Zhang (2001), and Brueckner (2004)


\[^4\] LCCs offer generally low fares in exchange for eliminating many traditional passenger services (such as first/business class and lounges) offered by full service airlines (FSA). The U.S. department of Transport (DOT) regularly reports financial and operational statistics for seven largest LCCs including Airtran Airways, American Trans Air, America West Airlines, Frontier, Jet Blue, Southwest, and Spirit.
large; (2) the distance between cities is very short; and (3) the number of cities is small. These predictions, of course, hardly explain the rapid expansion of Southwest and Jet Blue, which developed extensive networks with nationwide coverage. Running a PoP network does not necessarily mean that an airline has to manage each route independently. Two “adjacent” routes, such as routes linking the same two cities / regions out of alternative airports (for example, Los Angeles International (LAX) – Las Vegas (LAS) vs. Orange County (SNA) – Las Vegas (LAS), or Chicago O’Hare (ORD) – Newark (EWR) vs. Chicago O’Hare (ORD) – LaGuardia (LGA)), are imperfect substitutes to each other. Potential gains may be achieved by strategically managing these two routes together rather than treating them as unrelated markets. Also, by consolidating all route operations in one airport, an airline may achieve, at least partially, the cost reduction and market power effects identified for HS networks. These possible network effects can be important to explain the expansion of LCCs, and to understand the strategies adopted by legacy carriers (e.g., United and American airlines) to compete with LCCs. However, network effects in PoP networks have received little attention in the literature.

An airline’s profitability of operating different routes should be interdependent if there are significant network effects. Therefore, network effects can be empirically explored by examining the patterns of spatial dependence in profitability across different routes. Following the standard approach of modeling firm entry, we treat the observed entries of an airline as the indicator of the underlying profitability. We further classify the relationships among all routes into different categories of spatial dependence based on the distance between end-point airports. The dependence of the underlying profitability across routes can then be uncovered from the spatial entry patterns of the airline. We choose to study the entry decisions of Southwest, the largest and most successful LCC in the U.S..
data cover the period of 1990 to 2006, during which the airline experienced tremendous growth. $^5$

Significant spatial dependence has been identified in Southwest’s route entries, which suggests two major sources of network effects in PoP networks: airport/regional presence and market substitutability; network effects embedded in Southwest’s Point-to-Point network have many distinguishing features as compared to those identified in a typical Hub-and-Spoke network. Moreover, spatial entry patterns of Southwest changed significantly during our study period, calling for the need to control for airline size and network coverage when one attempts to identify alternative network effects.

Results of this paper bring fresh insights on the sources and mechanisms of airline network effects in general, as well as explain the pattern of aggressive network expansion of LCCs in particular. While previous studies (Dresner et al. 1996, Windle and Dresner 1999, Morrison 2001, and Hofer et al. 2007) suggest that the growth of LCCs has brought substantial welfare gains, this paper renders valuable empirical implications by identifying spatial dependence of airline entry decisions which is useful to evaluate the spill-over effects of LCC entry. This paper also extends the extensive literature on airline entry (for example, Berry 1992, Sinclair 1995, Boguslaski et al. 2004, Aguirregabiria and Ho 2007, Dixit and Chintagunta 2007), which has not modeled the spatial dependence in airline entry decisions explicitly. Finally, understanding spatial entry patterns of Southwest is helpful to understand market outcomes in airline industry (Goolsbee and Syverson (2008)).

This paper is organized as the follows: Section 2 reviews some of the emerging patterns observed in Southwest’s network development. Section 3 specifies the econometric model. Section 4

$^5$ From 1993 to 2004, the revenue and passenger volume of Southwest were almost tripled and the airline added service to 22 new airports (Goolsbee and Syverson, 2008)
describes the data and variables used in empirical analysis. Section 5 reports and evaluates the estimation results. The last section concludes the paper.

2. Some Observations on Southwest’s Network Development

There are two significant patterns in Southwest’s network expansion especially starting from late 1990’s: 1) once the airline started service at an airport, it entered multiple markets connecting the airport; 2) the airline expanded its network mainly by adding routes to connect existing airports rather than adding new airports. As shown in Table 1, Southwest increased the number of routes (airport-pairs) by more than 300%, from 81 routes in 1990 to 375 routes in 2006 while increasing the number of airports by only 100% during the same period. As a result, the average number of destinations out of an airport has increased from 5 in 1990 to 12 in 2006.

During its expansion, Southwest increased operations at several focus airports substantially. Table 2 lists the top five airports for Southwest over the years. Currently non-stop flights from these top five airports account for more than half of the Southwest system. Unlike hub airports serving large proportions (usually about 50%) of connecting passengers, Southwest’s focus airports serve mainly local passengers. Figure 1 reports the ratio of enplaned passenger (passengers getting on a flight at the focus airport) to onboard passenger (including both enplaned passenger and through passengers who stop by the focus airport before flying to final destination) at Southwest’s top five airports; in 2006, more than 80% of passengers at these airports are enplaned (local) ones. Connecting passengers at hub airports normally change airplane to fly to their destinations; through passengers of Southwest usually do not change airplane at focus airports.
Focus airports of Southwest can be quite close to each other. Examples include San Diego (SAN) and Los Angeles (LAX), Orlando (MCO) and Tampa (TPA), and Baltimore-Washington (BWI) and Philadelphia (PHL); distances between these airport pairs are all close to 100 miles. This is very different from HS system in which the minimum distance between two major hub airports in a successful dual-hub system is about 560 miles \(^6\) -- Northwest’s Minneapolis-St. Paul and Detroit. British Airways once attempted to share hub functions of London-Heathrow airport with Gatwick airport but soon realized that it is not beneficial (O’Connell, 2008).

Another noticeable feature in Southwest’s network configuration is that the airline tends to serve multiple airports in a metropolitan area/region simultaneously. For example, Southwest has operations at four airports in Los Angeles area, three airports in San Francisco bay area, and two airports in Washington, DC area; Southwest has also operations at close-by airports in Texas, Florida, and New England.

Above observations lead us to hypothesize that the PoP route system of Southwest is designed strategically to explore certain network effects. This hypothesis contradicts with the common view that a PoP system assembles a loosely connected collection of dense and short-haul markets, whose profitability are based on their own\(^7\).

3. The Econometric Model

\(^6\) AirNeth (2005).

\(^7\) This is the theoretical prediction from Hendricks et al. (1995 and 1999) and Lederer and Nambimadom (1998), as well as empirical finding from Boguslaski et al. (2004) on Southwest’s entry.
We use a *spatial probit model* to study the entries of Southwest Airlines on airport-pair markets; interdependence in profitability indicated by spatial entry patterns of the airline suggests existence of possible network effects. Airport-pair markets in our study are non-directional (i.e., MDW – JFK and JFK – MDW are considered as one market) because Southwest normally offers round-trip service between two airports. Spatial probit model was first introduced by McMillen (1992), and has been applied by several others on various problems (Case 1992, Marsh *et al.* 2000, Beron *et al.* 2003, Murdoch *et al.* 2003, and Coughlin *et al.* 2004).

In a spatial probit model, observations on the discrete dependent variable (route entries of Southwest in our case) are allowed to be interdependent in neighboring regions. Neighboring regions of an airport-pair market are defined by the distance from the end-point airports. In his study on airline competition, Morrison (2001) uses a zone with a radius of 75 miles to define “nearby” airports. He finds that the presence of Southwest at an airport affects airfares of markets out of nearby airports, particularly those “adjacent” routes – routes connecting nearby airports to nearby airports. He also finds that a zone with a radius of 75 miles provides the best fit compared to zones of 25, 50, 100, and 150-mile radius. We define the spatial relationship of airport-pair markets with similar terms used by Morrison. We consider interdependence among airports up to 150 miles apart, and allow such interdependence to be different for airports within 75 miles vis-a-vis airports more than 75 miles apart.

In particular, two airports are defined as *local* to each other if they are less than or equal to 75 miles apart. For example, Chicago O’Hare airport and Chicago Midway airport are local airports to each other, and the three airports in New York/New Jersey area are also local airports to each other.

---

8 MDW is the code for Chicago Midway Airport, while JFK stands for the John F. Kennedy International Airport in New York City.
Two airports are defined as *nearby* to each other if the distance between them is larger than 75 miles but less than or equal to 150 miles. For example, Los Angeles International Airport is a nearby airport to San Diego International Airport. Finally, two airports are *unrelated* if the distance between them is greater than 150 miles. The spatial relationship of any two airport-pair markets can therefore be classified into one of the 8 types as summarized in Table 3. For ease of notation, we use $m' \sim^j m$ to denote that market $m'$ is type $j$ related to market $m$ hereafter.

Let $y_m$ indicate the observed entry decision of Southwest on market $m$ in a year, and $M$ be the total number of airport-pair markets; $y_m$ equals to one if Southwest *served* the market in all four quarters of the year and equals to zero otherwise. This entry definition is different from the ones used in previous studies on airline entry, where an airline is said to enter a market if the firm *started* to serve the route in a year. While previous studies attempt to identify the factors affecting an airline’s decision to start new services, we are concerned about airlines’ strategy of network configuration, which should be better understood as a long-run decision. With our entry definition, entries observed in a year represent an airline’s network configuration, which is formed by all previous network decisions. If Southwest served a market from 1990 – 1994 but exited in 1995, we do not consider that the airline enters this market when we study its network configuration in 1995. A temporary entry/exit decision may reveal important factors affecting the airline’s decision to start/stop services in a market, but it is not suitable for our purpose to identify the network effects embedded in the long-run network configuration.

An airline’s entry decision in a market depends on the perceived post-entry profit, which is denoted by $\pi_m$. An airline only enters markets where profits are non-negative. Formally we have
Profits are assumed to be dependent on exogenous market factors and some factors are unobservable to econometrician; profit in one market depends also on profits in related markets because of possible network effects. We thus specify the profit function as the following

\[ \pi = XB + W\pi + \varepsilon \]  

(2)

where \( \pi = [\pi_1, \pi_2, \ldots, \pi_M] \) and \( X = [X_1, X_2, \ldots, X_M] \); \( X_m \) is a row vector of market factors related to market \( m \). \( W \) is a \( M \) by \( M \) spatial dependence matrix. The \( m \)th row and the \( n \)th column of the spatial dependence matrix is 0 if \( m = n \) or market \( m \) is unrelated with market \( n \); is \( \rho_j \) if \( n \sim^j m \), where \( \rho_j \) is the spatial lag coefficient. In total we have 8 such coefficients to estimate from data. The noise term \( \varepsilon = [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_M] \) represents the market factors which can be observed by the airline but not econometrician; the elements in \( \varepsilon \) are parameterized as mutually independent standard normal distributions.

Consider an example in which Southwest makes entry decisions on a network composed of the following four airport-pair markets: Los Angeles Airport – San Francisco Airport (market 1), Los Angeles Airport – Oakland Airport (market 2), San Diego Airport – Oakland Airport (market 3), and San Francisco Airport – Seattle Airport (market 4). The spatial dependence matrix of this network is

\[
W = \begin{bmatrix}
0 & \rho_1 & \rho_5 & \rho_3 \\
\rho_1 & 0 & \rho_2 & \rho_7 \\
\rho_5 & \rho_2 & 0 & \rho_7 \\
\rho_3 & \rho_7 & \rho_7 & 0
\end{bmatrix}
\]  

(3)
and thus the profit of entering market 1 is

$$\pi_1 = X_1B + \rho_1 \pi_2 + \rho_2 \pi_3 + \rho_3 \pi_4 + \varepsilon_1$$  \hspace{1cm} (4)$$

The profit of entering market 1 depends on the profits of entering other three markets, and this spatial dependence can be caused by possible network effects. Estimation results on the spatial lag coefficients reveal whether such network effects are material.

Equation (2) leads to the following reduced form

$$\pi = \tilde{X}B + \tilde{\varepsilon}$$  \hspace{1cm} (5)$$

where $$\tilde{X} = (I - W)^{-1} X$$, $$\tilde{\varepsilon} = (I - W)^{-1} \varepsilon$$ and I denotes a M by M identity matrix. The data likelihood of the airline’s entry decisions on the M markets is thus

$$L(B, \rho) = \prod_{m=1}^{M} \Phi \left( \frac{\tilde{x}_m B}{\sigma} \right) \left[ \frac{1 - \Phi \left( \frac{\tilde{x}_m B}{\sigma} \right)}{1 - \Phi \left( \frac{\tilde{x}_m B}{\sigma} \right)} \right]$$  \hspace{1cm} (6)$$

where $$\rho = [\rho_1, \rho_2, \ldots, \rho_8]$$; $$\tilde{x}_m$$ is the mth row of $$\tilde{X}$$; $$\sigma^2 = \sigma_m \sigma'_m$$ with $$\sigma_m$$ denoting the mth row of $$(I - W)^{-1}$$, and it is the function of spatial lag coefficients $$\rho$$; $$\Phi(\cdot)$$ is the c.d.f. of the standard normal distribution.

The model presented here explicitly accounts for the interdependence of profitability among different markets. This is the key feature of our model compared with previous airline entry studies. Earlier studies admit that airlines’ entry decisions can be largely driven by their network strategies; they explore the network effects by studying how an airline’s post-entry profit in a market is affected
by its pre-entry presence at end-point airports. Such studies assume that airlines have no overall strategy for network configuration and they *sequentially* optimize their route entries *conditional on* their previous decisions. Network of an airline at a time point is the result of all previous entry decisions driven by market factors and the purpose of exploring network effects; although entries are sequential, decisions of these entries are likely to be affected by common factors. Furthermore, network effects may exist among not only markets connecting the same airports but also markets connecting nearby/local airports. Profitability interdependence and network effects, if any, can be identified with our model specification by examining the estimation results of the spatial dependence matrix.

The parameters in the spatial probit model can be estimated by the Maximum Likelihood Estimator (MLE) which maximizes the log of the likelihood function in (6). The computation burden, however, is heavy because operating a large size spatial dependence matrix without convenient structure cannot be avoided in evaluating the likelihood function. The network considered in our analysis includes the top 3000 airport-pair routes in the US domestic market. The spatial dependence matrix is then 3000 by 3000, and it has no convenient structure such as block diagonal as in many other spatial studies. With MLE, we have to evaluate $(I - W)^{-1}$ in each iteration of the optimization process\(^9\), which makes the computation burden very heavy. In order to avoid such difficulty in computation, we adopt a Bayesian approach which is described in details in the appendix.

4. The data

---

\(^9\) Because the elements in $W$ is the spatial lag coefficients.
We study the entry patterns of Southwest Airlines from year 1990 to 2006, during which the airline expanded its network substantially. The study is conducted over the largest 3000 non-stop airport-pair markets, which accounted for more than 90% of all domestic traffic as of the first quarter in year 2000. We divide the 17 years into three time windows, namely 1990-1995, 1996-2000, and 2001-2006. Entries in the last year of each period are used for analysis. The multiple time windows allow us to track the evolvement of the airline’s network configuration; the 5-6 year interval of each period is to incorporate the adjustment process of the airline’s network development. The dependent variable of the model for each of the three periods is the entry dummy defined in (1), which takes the value of one in a market if Southwest operates at least one non-stop flight per week in the last year of the period. The entry information is compiled with the Department of Transport (DOT) T100 data for airline scheduled departure. In our specification, Southwest’s entry, or the post-entry profitability in an airport-pair market, depends on the following factors:

- **Distance**: Distance in miles between the origin and destination airports in each potential market considered. It is well known that route distance affect potential demands for air travel in a market as well as airlines’ overall network configuration.

- **Population**: Mean population for the catchment areas of the two airports. If an airport is located in a Metropolitan Statistical Area (MSA)\(^{10}\), then the MSA region is used as the airport’s catchment area. Otherwise the catchment area is defined as the county in which the airport is located. Population data are from the Regional Economic Information System (REIS) maintained by the Bureau of Economic Analysis (BEA).\(^{11}\)

---

\(^{10}\) As defined by the U.S. Office of Management and Budget, released in June 2003 with revisions released in February 2004 and December 2005.

\(^{11}\) Regional Economic Information System, 1969-2005 version.
o **Income**: Mean income per capita for the two airports’ catchment areas. Income per capita data are from the REIS system maintained by the BEA.

o **Maximal airport HHI**: Based on the number of fared passengers carried by each airline, HHI index is calculated for each airport. A higher HHI indicates a more concentrated market structure at an airport. If the airport HHI index is high for one of the two airports in a market, low cost carriers may face higher entry costs. We therefore incorporate the maximum of airport HHI on a market in our empirical specification. Also, in order to avoid potential endogenous problem, we use the pre-entry airport HHI which is calculated from the 1990 data. Fared passenger data are compiled from the DB1A database maintained by the DOT. The database contains 10% sample of all US domestic origin and destination tickets.\(^{12}\)

o **Maximum airport volume**: From the fared passenger volume data, we compile the passenger volume out of each airport. Low cost carriers such as Southwest Airlines traditionally avoided large airports with congestion. The maximum airport volume of the two airports is thus used in order to control for such effects. Again, in order to avoid potential endogenous problem, we use pre-entry airport volume dates which are calculated from the 1990 data.

o **Hub Market and Double Hub Market**: Those are dummy variables for a city-pair market if at least one airport is the hub of a major Full Service Airline (FSA). If both airports in a route are hubs of some major FSAs, it is defined as a Double Hub Market. Our hub definition follows Brueckner (2002).\(^{13}\)

\(^{12}\) We purchased the data from Database Product Inc., a reseller of the DOT airline data.

\(^{13}\) Brueckner (2002) defined following hubs in the U.S. markets: Chicago-O’Hare (American. United), Cleveland (Continental), Newark (Continental), Atlanta (Delta), San Francisco (United), Dallas-Ft. Worth (American), Philadelphia (US Airways), Phoenix (America West), Detroit (Northwest), St. Louis (American), Houston (Continental), Washington-Dulles (United), Minneapolis-St. Paul (Northwest), Cincinnati (Delta), and Miami (American).
- **Vacation**: This is a dummy variable for tourism routes. An airport-pair market is defined as a vocational route if at least one of the airports is located in Florida or Nevada. Such a definition has been used in Dresner et al. (1996) and Windle and Drenser (1999).

- **Medium Market and Large Market**: These are dummy variables used to capture possible non-linear effects of market potential on Southwest’s entry decisions. An airport-pair market is classified as “medium” if the yearly fared passenger volume is between 35,000 to 80,000 in year 1990. If there are more than 80,000 yearly fared passengers in 1990, the market is defined as “large”.

Table 4 summarizes these variables and their summary statistics.

### 5. Results

Estimation results are presented in Table 5. We present the posterior mean and posterior probability of being positive (numbers in parentheses) for each of the coefficients. For example, the posterior mean of medium market size coefficient in the model in 2006 is 0.7533, and the posterior probability of this coefficient being positive is 100%. This suggests that the chance of this coefficient being negative is zero. In another example, the posterior mean of income coefficient in the model in 2006 is -0.0750 and the posterior probability of this coefficient being positive is zero. This suggests that the chance of this coefficient being negative is 100%; Southwest’s post-entry profitability decreases when the mean income of the two cities linking the route gets increased.

---

14 Morrison (2001) proposed a similar definition (“Sunbelt airports”) for tourism routes. Previous studies have found that the two definitions have essentially the same prediction power.
The first panel of Table 5 shows the impacts of market factors on Southwest’s network configuration. Some market factors affect the airline’s network choice consistently. First, as expected from theories, Southwest tends to enter dense markets throughout all the three periods. Second, Southwest tends to enter markets with low average per capita income. This is probably due to the fact that the airline prefers to use secondary airports to serve price sensitive consumers. Finally, Southwest tries to avoid hub airports or airports which are highly concentrated, as indicated by the results of coefficients for airport HHI index and hub dummy.

The effects of some market factors on Southwest’s network development evolved over time. For example, the magnitude of the negative impact of airport HHI index on Southwest entry decision kept decreasing and the negative impact of distance on Southwest’s entry decision diminished over the years. While there may be more than one driving factors for such changes, estimation results on spatial dependence coefficients in our model reveal that network effects embedded in Southwest’s network development cause the changes. We examine such effects in details below.

*Estimation results on spatial dependence coefficients*

The second panel of Table 5 presents the results on the spatial dependence coefficients. Among the eight spatial dependence coefficients, type 1 – 3 coefficients are positive in general (except type 1 coefficient in 1995); type 6 and type 8 coefficients are positive in 1995 but became negative along with type 5 and type 7 coefficients in 2006. These results indicate network externalities existed in Southwest’s PoP route system and suggest network effects from two sources – *airport/regional presence* and *market substitutability*.

---

15 In the 2006 model (the posterior mean of the distance coefficient is -0.0058 and the posterior probability of the coefficient being positive is 0.19
Let’s first look at estimation results for type 1-3 coefficients, all of which capture spatial dependence in profitability of markets connecting the same airport. The results can be summarized as follows. First, in Southwest’s route system, profitability of markets out of the same airport is in general positively dependent. Second, such positive dependence has become stronger and stronger over the years. Third, positive dependence among type 1 and type 2 markets is greater than the one among type 3 markets. These results imply that Southwest enjoys positive network externalities through large scale of operations at focus airports; network externalities are further enhanced when markets out of the same focus airport are imperfect substitutes (type 1 and type 2). These combined effects explain the observed patterns in Southwest’s network developments over the years, especially expansions starting from late 1990’s: to launch multiple routes out of a new added airport; to increase scale of operations at focus airports substantially; to connect multiple airports in a metropolitan area/region with a focus airport.

The network effects in Southwest’s PoP route system captured by type 1-3 coefficients are similar to those achieved through hubbing in HS route systems. Nevertheless, the sources and mechanisms of these network effects are different. As mentioned briefly in the introduction, in a HS route system an airline’s profitability in spoke markets out of the same hub airport is positively related to the airline’s scale of operations at the hub because: a) traffic density in spoke markets can be increased through hubbing such that production costs at each of the spoke markets are reduced with economies of traffic density; b) flight frequency on spoke markets can be increased through hubbing such that service quality is improved; c) different spoke markets connecting the same hub airport are complements to each other such that the hub operator can compete more strategically.\textsuperscript{16} The recent

\textsuperscript{16} Hendricks et al. (1997) show that the hub operator partially offset financial loss in one spoke market by adjusting decisions on other spoke markets
study by Aguirrebabiria and Ho (2008) suggests two additional explanations. First, the hub operator’s entry cost into a spoke market connecting the hub is lower if operator has larger scale of operations at the hub airport. Second, larger scale of operations at the hub can deter competitors’ entries into spoke markets more effectively.

The percentages of connecting passengers at Southwest’s focus airports are quite low compared to HS network carriers. Therefore, among the explanations on network effects through hubbing, only the two proposed in Aguirrebabiria and Ho (2008) are relevant in explaining the positive dependence in profitability of markets out of the same airport in Southwest’s route system. There is another important distinction between Southwest’s PoP system and HS systems. In Southwest’s PoP system, the positive dependence in profitability of markets out of the same airport is strengthened by substitutability of markets, whereas such dependence in HS systems comes partly from the complementarity of spoke markets. Serving substitute markets together can be beneficial for Southwest. First of all, like a hub operator, Southwest can affect competition outcomes in one market by strategically adjusting decisions in other markets\(^\text{17}\). Moreover, by serving substitute markets together, Southwest can deter potential entries into the origin-destination region, or compete with incumbents in the region more effectively. These network effects should be particularly evident for a group of routes which are highly substitutable to each other, such as the type 1 or 2 markets in our analysis which have substantially overlapping catchment areas around endpoint airports. This is the case for routes out of Los Angeles region, where four airports (LAX, SNA, ONT, and BUR) are “local” to each other. As of 1990, American West, another major LCC in the U.S., had significant

\(^{17}\) Discussions on a firm’s incentives and mechanisms to provide substitute products can be found in the literatures of product selection, analysis of multiproduct firms and product line rivalry. A review of these models and their assumptions would be useful, it is nevertheless beyond the goals this study attempts to achieve.
presence in all four airports and was dominant/major carrier in markets connecting these four airports to Las Vegas and Phoenix. The same strategy of network development has been used by Southwest; the airline started to enter all these markets from 1990’s, gaining market shares from American West in all routes.18

Southwest has focus airports which are not in the same metropolitan area yet close to each other. Because of this network strategy we observe a large number of type 2 markets in the carrier’s route system. Such a network design allows the carrier to exploit combined network effects from airport/regional presence and market substitutability. This is not the case for a typical HS system, where much of the network effects come from increased traffic density, market power at hub airports and complementarity among spoke routes. If two hub airports were close to each other, traffic volume would have to be shared among competing markets, leading to reduced traffic density, reduced market power at hub airports and higher operation costs. This explains why British Airways failed to run both London-Heathrow airport and Gatwick airport as its hub airports.

Unlike type 1-2 markets, type 3 markets are not substitutes to each other. Southwest’s network strategy to launch multiple routes from the same airport creates lots of type 3 markets in its route system. Relationships among these type 3 markets can be similar to those among spoke markets in a HS system if Southwest provides connecting services at focus airports. Southwest does provide

18 Brander and Eaton (1984) studied product line rivalry where firms offer substitute products in the market. They found that “sequential decisions on product type and output can naturally give rise to equilibria in which a single firm monopolizes close substitutes. Such outcomes hold only for a certain levels of demand and might, therefore, be observed only over some portion of the life cycle of the industry”. Studying a sequential game in which two duopoly firms choose to produce substitutable products, they show that it is possible that the two firms segment the market. However, “as growth occurred in the market and the game was repeated, market segmentation would be replaced by market overlapping”. The pattern of (American West) Monopoly – (American West and Southwest) Segmentation during the process of Southwest’s entry – Market overlapping (after Southwest’s entry in all routes) appeared to be consistent with such theoretical results on substitute products.
connecting services to combine traffic volume, as found by Fu et al. (2007). However, the mechanism and extent are very different from HS carriers. To provide extensive connection services at hub airports, HS carriers need to coordinate/schedule flights to arrive and depart at almost the same time. This is complex and demands expensive peak time capacities. Since Southwest mainly uses secondary airports with limited capacity, such a strategy may not even be feasible. Instead, Southwest lets connecting (“through”) passengers to stay in the aircraft while local passengers get onboard, in order to maintain a short turn-around time at focus airports. Such an operation allows the carrier to combine traffic in its PoP route system to some extent, without incurring the congestion and operations costs of a HS system. Although connecting services at focus airports explain part of the positive dependence in profitability of type 3 markets, it is unlikely to be the major cause of network effects given the low shares of connecting passengers.

Estimation results on other spatial dependence coefficients confirm the existence of network externalities from regional presence and market substitutability. Table 5 shows that the profitabilities of markets connecting nearby airports to nearby airports (type 6) and of markets connecting nearby airports to unrelated airports (type 8) are both positively dependent in early 1990’s. However, in the 2000’s, type 5 – 8 spatial dependence coefficients are all negative. Such a change over time suggests that Southwest has consolidated its route system to combine the two sources of network effects, namely airport/regional presence and substitutability of markets. Figure 2 gives an illustrative example: rather than serving multiple airports in a region (local or nearby airports), Southwest tends to serve multiple markets from one airport in order to strengthen network effects from airport presence. When serving substitute markets connecting the same endpoint airports (type 1 and type 2 markets), Southwest can combine the two sources of network effects in its route system. With such a network strategy, type 5 – 8 spatial dependence coefficients become negative.
Comparison between airlines

Spatial entry patterns of Southwest suggest existence of network effects in the carrier’s PoP route system, with the two main sources being airport/regional presence and market substitutability. These findings, together with the conclusions obtained in previous studies on HS networks, suggest that the network effects embedded in Southwest’s PoP system have many distinguishing features compared to those identified in a typical HS system. One way to test such distinguishing features is to compare the entry patterns between Southwest and a HS carrier. An ideal “control group” to Southwest for such a comparison is another major low cost carrier in US domestic market running HS network. AirTran meets all these requirements and hence, is chosen for such a comparison. Due to the lack of enough entry observations for AirTran in the 1990’s, the comparison is only made for recent years (up to 2006).

Table 6 presents results of the comparison. It can be seen that both airlines prefer dense routes linking tourism destinations. However, these airlines’ have different marketing and operation strategies: while Southwest has always been a cost leader offering no-frill services, AirTran provides onboard amenities, such as business class and satellite radio channels, comparable to many full service airlines such as United and American Airlines. This probably explains the positive and high posterior probability coefficient of income for AirTran. The carrier provides services out of major hubs at Atlanta and Orlando. This has been reflected in our estimates: the coefficients for Maximal airport volume dummy and Hub market dummy are both positive with high posterior probability.

Network effects embedded in the two carriers’ route systems, as revealed by the estimation results of spatial dependence coefficients, are very different. In AirTran’s route system, profitability of markets connecting the same airport to unrelated ones is positively dependent, but profitability of
markets connecting the same airport to local/nearby ones is negatively dependent. As discussed previously, connecting complementary spoke routes via a hub airport can increase the traffic density (thereby reduce cost) of spoke markets. Therefore, type 3 spatial dependence coefficient is positive in AirTran’s model. Meanwhile, to achieve such network effects through hubbing, spoke markets cannot be substitutes to each other. This explains the negative type 1 and type 2 spatial dependence coefficients in AirTran’s model. These comparison results confirm the distinguishing features identified for PoP networks vis-à-vis HS networks.

6. Conclusion

Clear spatial dependence patterns have been identified for Southwest’s entries in the U.S. domestic market throughout the sample period from 1990 to 2006. The carrier’s spatial entry pattern reveals existence of network effects in its PoP route system. Such network effects come from two main sources: (1) airport and regional presence, and (2) substitutability of markets. Over the years, Southwest has re-configured much of its route network in order to fully exploit the combined network effects. With these network effects strengthened, the importance of other factors affecting Southwest’s entry decisions such as market size, distance, and market concentration as measured by HHI index, had been diminishing over time.

Airport presence as a source of network effects has been well documented in previous studies on HS networks. Our study suggests that the mechanisms of generating network effects through airport presence are different in PoP and HS networks. In a HS system, airport presence is strengthened by coordinating extensive connection services among a large number of flights at hub airports. This leads to increased traffic volume and flight frequency in spoke markets, and makes spoke markets complement to each other. In a PoP system as operated by Southwest, airport presence is strengthened
by (1) combining “through” traffic with local passengers at focus airports. This allows the carrier to achieve limited traffic aggregation without incurring the costs related to running a hub; (2) serving substitute markets out of the same airports. This enhances the carrier’s market power in an origin-destination region, and allows Southwest to configure multiple focus airports close to each other. This hasn’t been possible for HS carriers, since locating two hub airports close to each other would reduce traffic volumes at hubs and spokes markets. To fully exploit the two sources of network effects jointly, i.e., airport / regional presence and market substitutability, Southwest had optimized its network over our sample period.

Our exploratory study reveals possible network effects in PoP route systems. The findings are helpful in understanding the rapid expansions of low-cost carriers such as Southwest, and outcomes of competition in the U.S. domestic airline market. However, to complement the exploratory findings in this paper, detailed investigation on the specific sources and mechanisms of the network effects would be very valuable. This is left as a future extension to the current study.
Table 1. Summary Statistics for Southwest’s Network in Selected Years

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of markets served (airport-pair routes)</td>
<td>81</td>
<td>152</td>
<td>295</td>
<td>375</td>
</tr>
<tr>
<td>Number of airports served</td>
<td>31</td>
<td>46</td>
<td>56</td>
<td>62</td>
</tr>
<tr>
<td>Average number of markets connecting one airport</td>
<td>5</td>
<td>7</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Number of airports with less than or equal to 3 routes</td>
<td>12</td>
<td>14</td>
<td>6</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 2. Five Most Concentrated Airports for Southwest Airlines in Selected Years

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Airport Rank</td>
<td>Airport</td>
<td>Route #</td>
<td>Airport</td>
<td>Route #</td>
</tr>
<tr>
<td>1</td>
<td>PHX</td>
<td>16</td>
<td>PHX</td>
<td>23</td>
</tr>
<tr>
<td>2</td>
<td>HOU</td>
<td>14</td>
<td>LAS</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>DAL</td>
<td>13</td>
<td>STL</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>ELP</td>
<td>12</td>
<td>HOU</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>ABQ</td>
<td>10</td>
<td>DAL</td>
<td>13</td>
</tr>
</tbody>
</table>

Notes: The airports included in the above table include Las Vegas (LAS), Chicago Midway (MDW), Phoenix (PHX), Baltimore-Washington International (BWI), Orlando (MCO), Nashville (BNA), Houston Hobby (HOU), Dallas Love Field (DAL), St. Louis (STL), El Paso International (ELP), Albuquerque International (ABQ).
Table 3. Spatial relationship of airport-pair markets

<table>
<thead>
<tr>
<th></th>
<th>Same</th>
<th>Local</th>
<th>Nearby</th>
<th>Unrelated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same</td>
<td>N.A.</td>
<td>Type 1</td>
<td>Type 2</td>
<td>Type 3</td>
</tr>
<tr>
<td>Local</td>
<td>Type 1</td>
<td>Type 4</td>
<td>Type 5</td>
<td>Type 7</td>
</tr>
<tr>
<td>Nearby</td>
<td>Type 2</td>
<td>Type 5</td>
<td>Type 6</td>
<td>Type 8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Example: LAX – SFO and SAN – SMF)</td>
<td>(Example: LAX – SFO and SAN – SEA)</td>
</tr>
<tr>
<td>Unrelated</td>
<td>Type 3</td>
<td>Type 7</td>
<td>Type 8</td>
<td>No relationship</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Example: LAX – SEA and SFO – BOS)</td>
</tr>
</tbody>
</table>

Note: LAX: Los Angeles Airport; SFO: San Francisco Airport; OAK: Oakland Airport; SAN: San Diego Airport; SMF: Sacramento Airport; SEA: Seattle Airport; SNA: Santa Ana Airport (Orange County, California); BOS: Boston Airport
<table>
<thead>
<tr>
<th>Variables</th>
<th>Description</th>
<th>Summary Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medium market size dummy</td>
<td>1 if the number of fared OD passengers in 1990 is greater than 35000 per year but less than or equal to 80000 per year</td>
<td>24%</td>
</tr>
<tr>
<td>Large market size dummy</td>
<td>1 if the number of fared OD passengers in 1990 is greater than 80000 per year</td>
<td>26%</td>
</tr>
<tr>
<td>Distance in miles</td>
<td>Distance between the two airports of a market</td>
<td>1027 (642)</td>
</tr>
<tr>
<td>Population in million</td>
<td>Mean population of the two MSAs or counties in which the two airports are located</td>
<td>3.00 (2.53)</td>
</tr>
<tr>
<td>Income in thousand US $</td>
<td>Mean income per capita of the two MSAs or counties in which the two airports are located</td>
<td>20.64 (21.31)</td>
</tr>
<tr>
<td>Maximal airport HHI</td>
<td>The maximum of Herfindahl-Hirschman Index of the two airports in 1990</td>
<td>0.26 (0.01)</td>
</tr>
<tr>
<td>Maximal airport volume in million</td>
<td>The maximum of passenger volume of the two airports in 1990</td>
<td>3.2 (1.6)</td>
</tr>
<tr>
<td>Hub market dummy</td>
<td>1 if at least one of the two airports is the hub of some major full service airlines</td>
<td>44%</td>
</tr>
<tr>
<td>Double hub market dummy</td>
<td>1 if both the two airports are hubs of some major full service airlines</td>
<td>4%</td>
</tr>
<tr>
<td>Vacation dummy</td>
<td>1 if at least one of the two airports is located in Florida or Nevada</td>
<td>26%</td>
</tr>
</tbody>
</table>

Note: We report sample means and sample standard deviations (numbers in parentheses) for continuous variables, and fractions of the sample for dummy variables.
Table 5. Estimation results for Southwest airline (Dependent Variable: Entry Dummy)

<table>
<thead>
<tr>
<th>Variables</th>
<th>1995</th>
<th>2000</th>
<th>2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>3.6695 (0.99)</td>
<td>3.3202 (1.00)</td>
<td>1.8357 (0.99)</td>
</tr>
<tr>
<td>Medium market size dummy</td>
<td>1.0533 (1.00)</td>
<td>0.8613 (1.00)</td>
<td>0.7533 (1.00)</td>
</tr>
<tr>
<td>Large market size dummy</td>
<td>2.2311 (1.00)</td>
<td>1.7812 (1.00)</td>
<td>1.6597 (1.00)</td>
</tr>
<tr>
<td>Distance (000’s miles)</td>
<td>-0.0935 (0.00)</td>
<td>-0.0239 (0.00)</td>
<td>-0.0058 (0.19)</td>
</tr>
<tr>
<td>Population (million)</td>
<td>0.0137 (0.66)</td>
<td>-0.0295 (0.11)</td>
<td>-0.0364 (0.41)</td>
</tr>
<tr>
<td>Income (000’s US $)</td>
<td>-0.0248 (0.20)</td>
<td>-0.0763 (0.00)</td>
<td>-0.0750 (0.00)</td>
</tr>
<tr>
<td>Maximal airport HHI</td>
<td>-11.835 (0.00)</td>
<td>-8.2741 (0.00)</td>
<td>-3.1936 (0.09)</td>
</tr>
<tr>
<td>Maximal airport volume (million)</td>
<td>-0.0362 (0.20)</td>
<td>-0.0118 (0.35)</td>
<td>-0.0056 (0.42)</td>
</tr>
<tr>
<td>Hub market dummy</td>
<td>-0.1596 (0.07)</td>
<td>-0.2069 (0.01)</td>
<td>-0.1203 (0.06)</td>
</tr>
<tr>
<td>Double hub market dummy</td>
<td>0.1988 (0.77)</td>
<td>0.1848 (0.77)</td>
<td>0.1763 (0.78)</td>
</tr>
<tr>
<td>Vacation dummy</td>
<td>-0.0237 (0.43)</td>
<td>0.2348 (1.00)</td>
<td>0.2544 (1.00)</td>
</tr>
</tbody>
</table>

**Spatial Dependence Coefficients**

<table>
<thead>
<tr>
<th>Type</th>
<th>1995</th>
<th>2000</th>
<th>2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 1 (same - local)</td>
<td>-0.0071 (0.38)</td>
<td>0.0338 (0.95)</td>
<td>0.0523 (0.98)</td>
</tr>
<tr>
<td>Type 2 (same - nearby)</td>
<td>0.0268 (0.98)</td>
<td>0.0493 (1.00)</td>
<td>0.0606 (1.00)</td>
</tr>
<tr>
<td>Type 3 (same - unrelated)</td>
<td>0.0053 (1.00)</td>
<td>0.0057 (1.00)</td>
<td>0.0061 (1.00)</td>
</tr>
<tr>
<td>Type 4 (local - local)</td>
<td>0.0094 (0.66)</td>
<td>-0.0226 (0.23)</td>
<td>0.0036 (0.56)</td>
</tr>
<tr>
<td>Type 5 (local - nearby)</td>
<td>-0.2262 (0.62)</td>
<td>-0.0159 (0.10)</td>
<td>-0.0195 (0.02)</td>
</tr>
<tr>
<td>Type 6 (nearby - nearby)</td>
<td>0.0018 (0.99)</td>
<td>-0.0004 (0.28)</td>
<td>-0.0008 (0.04)</td>
</tr>
<tr>
<td>Type 7 (local - unrelated)</td>
<td>-0.0059 (0.32)</td>
<td>-0.0116 (0.14)</td>
<td>-0.0113 (0.10)</td>
</tr>
<tr>
<td>Type 8 (nearby - unrelated)</td>
<td>0.0013 (1.00)</td>
<td>0.0002 (0.73)</td>
<td>-0.0004 (0.09)</td>
</tr>
</tbody>
</table>

Note: We present posterior means of coefficients, and the numbers in parentheses are $p(\beta > 0|Data)$, that is, the posterior probabilities that the coefficients are positive.
<table>
<thead>
<tr>
<th>Variables</th>
<th>Southwest 2006</th>
<th>AirTran 2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>1.8357 (0.99)</td>
<td>-2.6142 (0.01)</td>
</tr>
<tr>
<td>Medium market size dummy</td>
<td>0.7533 (1.00)</td>
<td>0.1472 (0.85)</td>
</tr>
<tr>
<td>Large market size dummy</td>
<td>1.6597 (1.00)</td>
<td>1.0179 (1.00)</td>
</tr>
<tr>
<td>Distance (000’s miles)</td>
<td>-0.0058 (0.19)</td>
<td>-0.0348 (0.00)</td>
</tr>
<tr>
<td>Population (million)</td>
<td>-0.0364 (0.41)</td>
<td>-0.0537 (0.04)</td>
</tr>
<tr>
<td>Income (000’s US $)</td>
<td>-0.0750 (0.00)</td>
<td>0.0538 (0.97)</td>
</tr>
<tr>
<td>Maximal airport HHI</td>
<td>-3.1936 (0.09)</td>
<td>1.6665 (0.69)</td>
</tr>
<tr>
<td>Maximal airport volume (million)</td>
<td>-0.0056 (0.42)</td>
<td>0.2651 (1.00)</td>
</tr>
<tr>
<td>Hub market dummy</td>
<td>-0.1203 (0.06)</td>
<td>0.2253 (0.98)</td>
</tr>
<tr>
<td>Double hub market dummy</td>
<td>0.1763 (0.78)</td>
<td>0.5489 (0.99)</td>
</tr>
<tr>
<td>Vacation dummy</td>
<td>0.2544 (1.00)</td>
<td>0.7075 (1.00)</td>
</tr>
</tbody>
</table>

**Spatial Dependence Coefficients**

<table>
<thead>
<tr>
<th>Type</th>
<th>Southwest 2006</th>
<th>AirTran 2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 1 (same - local)</td>
<td>0.0523 (0.98)</td>
<td>-0.0417 (0.05)</td>
</tr>
<tr>
<td>Type 2 (same - nearby)</td>
<td>0.0606 (1.00)</td>
<td>-0.0147 (0.14)</td>
</tr>
<tr>
<td>Type 3 (same - unrelated)</td>
<td>0.0061 (1.00)</td>
<td>0.0067 (1.00)</td>
</tr>
<tr>
<td>Type 4 (local - local)</td>
<td>0.0036 (0.56)</td>
<td>0.0243 (0.77)</td>
</tr>
<tr>
<td>Type 5 (local - nearby)</td>
<td>-0.0195 (0.02)</td>
<td>-0.0081 (0.29)</td>
</tr>
<tr>
<td>Type 6 (nearby - nearby)</td>
<td>-0.0008 (0.04)</td>
<td>0.0027 (1.00)</td>
</tr>
<tr>
<td>Type 7 (local - unrelated)</td>
<td>-0.0113 (0.10)</td>
<td>-0.0172 (0.09)</td>
</tr>
<tr>
<td>Type 8 (nearby - unrelated)</td>
<td>-0.0004 (0.09)</td>
<td>0.0011 (1.00)</td>
</tr>
</tbody>
</table>

Note: We present posterior means of coefficients, and the numbers in parentheses are $p(\beta > 0|Data)$, that is, the posterior probabilities that the coefficients are positive.
Figure 1. Enplaned Passenger to Onboard Passenger Ratio for the top 5 focus airports in 2006

Note: Enplaned passengers refer to passengers who get on (enplane) a flight at an airport. Onboard passengers refer to all passengers in a flight at the airport, which include both enplaned passenger, and “through” passengers who fly to the focus airport, remain in the aircraft and fly to the next destination.
Figure 2. An illustration for Southwest's network strategy to combine sources of network effects. A and F are two closely related (local or nearby) airports. D and E are also two closely related airports. Airport A is a focus airport of Southwest. Instead of entering market EF, Southwest enters market AE in order to strengthen network effects from both airport presence and market substitutability.
Appendix. A Bayesian Approach for Estimating the Spatial Probit Model

As in Albert and Chib (1993), we treat latent profits $\pi_i$ as parameters and then work with an augmented posterior of the form

$$p(B_i, \rho_i, \pi_i | Data) \propto p(B_i, \rho_i) \times \prod_{m=1}^{M} \left[ \phi(\pi_{im}; x_{im}B_i + z_{im}\rho_i, \|\pi_{im} \geq 0) \times I(y_{im} = 1) + \right.$$ \left. \phi(\pi_{im}; x_{im}B_i + z_{im}\rho_i, \|\pi_{im} < 0) \times I(y_{im} = 0) \right]$$

(A.1)

where $p(B_i, \rho_i)$ is the prior of the parameters. $\phi(x; a, b | x \geq 0)$ denotes a univariate truncated normal density in which the normal distribution of $x$ (with the mean of $a$ and the variance of $b$) is truncated below zero (truncated above zero for $\phi(x; a, b | x < 0)$); $z_{im}$ is a 1 by 8 vector in which the $j$th element is $\sum_{m'} \pi_{im} \times I(m' \sim j) m$; $I(\cdot)$ represents the indicator function. In estimation, the prior of parameters is chosen as

$$p(B_i, \rho_i) = N(B_i; B_0, V_B) \times N(\rho_i; \rho_0, V_{\rho})$$

(A.2)

where $N(B_i; B_0, V_B)$ is the multivariate normal density with the mean vector of $B_0$ and the variance-covariance matrix of $V_B$ (similar for $N(\rho_i; \rho_0, V_{\rho})$). $B_0, V_B, \rho_0,$ and $V_{\rho}$ are the hyperparameters which are chosen to have a very diffuse prior in estimation.

The posterior density in (A.1) is complicated such that it is not possible for us to derive analytical properties of it. Alternatively, we use the Monte-Carlo simulation to simulate random draws from the posterior and the empirical moments of the random draws will be used to summarize estimation results. Although the joint posterior in (A.1) is not known explicitly, the conditional
distributions \( p\left(\pi_{im}\mid\pi_{im'},B_i,\rho_i,Data\right) \) for each \( m = 1,\ldots,M \), \( p(B_i\mid\rho_i,\pi_i,Data) \), and \( p(\rho_i\mid B_i,\pi_i,Data) \) are all standard distributions presented as follows:

1. For each \( m \), \( p\left(\pi_{im}\mid\pi_{im'},B_i,\rho_i,Data\right) \approx \begin{cases} \phi\left(\pi_{im};\gamma_iB_i + z_{im}\rho_i,1\mid\pi_{im} \geq 0\right) & \text{if } y_{im} = 1 \\ \phi\left(\pi_{im};\gamma_iB_i + z_{im}\rho_i,1\mid\pi_{im} \geq 0\right) & \text{if } y_{im} = 0 \end{cases} \)

2. \( p(B_i\mid\rho_i,\pi_i,Data) \approx N\left(B_i; Dd, D\right) \), where \( D = \left(X_i'X_i + V^{-1}_\beta\right)^{-1} \), \( d = \left(X_i'\tilde{\pi}_i + V^{-1}_\beta B_0\right) \), and 
\( \tilde{\pi}_i = \pi_i - Z_i\rho \), with \( Z_i = [z_{i1}, z_{i2}, \ldots, z_{iM}]' \).

3. \( p(\rho_i\mid B_i,\pi_i,Data) \approx N\left(\rho_i; Dd, D\right) \), where \( D = \left(Z_i'Z_i + V^{-1}_\rho\right)^{-1} \), \( d = \left(Z_i'\tilde{\pi}_i + V^{-1}_\rho \rho_0\right) \), and 
\( \tilde{\pi}_i = \pi_i - X_iB_i \)

A Gibbs sampling, which is to generate an instance from these conditional distributions (conditional on the current values of other variables) in turn, is therefore applicable for our purpose. As shown by Gelman et al. (2004), the sequence from the Gibbs sampling comprises a Markov Chain and the stationary distribution of the Markov Chain is just the joint posterior in (A.1).
References


35